Filtering non-stationary geophysical data with orthogonal wavelets

F. Moreau, D. Gibert, Ginette Saracco

GEOPHYSICAL RESEARCH LETTERS, VOL. 23, NO. 4, PAGES 407-410, FEBRUARY 15, 1996

Filtering non-stationary geophysical data with orthogonal wavelets

Frédérique Moreau, Dominique Gibert, and Ginette Saracco

Géosciences Rennes - CNRS/INSU, Université de Rennes 1, France.

Abstract. A filtering method based on both orthogonal wavelet decomposition and chi-squared statistics is proposed to clean non-stationary signals embedded in a gaussian white noise. An application to a time series of thermistance data recorded in an underground quarry illustrates the interest of the technique.

Introduction

The wavelet theory (*Grossmann and Morlet* [1984]; see *Meyer* [1990], *Daubechies* [1992], and *Holschneider* [1995] for reviews) constitutes a powerful framework to process and analyze non-stationary geophysical signals [*Foufoula-Georgiou and Kumar*, 1994]. In particular, the important problem of signal denoising has recently been addressed by means of both continuous [e.g. *Mallat and Hwang*, 1992] and orthogonal wavelet decompositions [*Donoho and Johnstone*, 1994; *Saito*, 1994]. Denoising needs to distinguish the noise from the signal and, depending on the particular models assumed for these components, distinct algorithms may be proposed. In the present study we address the particular issue of cleaning signals embedded in gaussian white noise through orthogonal wavelet decomposition. We propose a special-purpose filtering criterion based on a Chi-Square Thresholding (CST), and compare its performance to those of two general-purpose and popular threshold criteria: the Akaike's Information Criterion (AIC) [*Akaike*, 1965] and the Minimum Description Length (MDL) [*Rissanen*, 1978; *Wax and Kailath*, 1985]. Examples with synthetic tests and real geophysical data are given.

Denoising Signals with Wavelets

An orthogonal basis reads $\left\{2^{-m/2}\psi\left(2^{-m}t-n\right)\right\}$ with $(m,n)\in\mathbb{Z}^2$ where the analyzing wavelet $\psi(x)$ is an oscillating function localized around the origin. The wavelet coefficients of a discrete signal may be efficiently computed via a pyramidal algorithm *Mallat* [1989] and provide a way to examine the information content of the original signal in the time-scale half-plane. If the input signal s counts $K=2^N$ values, the first 2^{N-1} wavelet coefficients correspond to the finest scale available and fixed by the sampling interval, the next 2^{N-2} coefficients are for the immediately upper scale (i.e. twice the finest scale), and so on until the last coefficient which corresponds to the largest scale available (i.e. the length of the signal). A filtered signal is obtained by performing an inverse wavelet transform with a subset of the initial K wavelet coefficients.

Filtering Criteria

Since each of the K coefficients may be either rejected or retained, the set $\mathcal{A}=\{s_l\}$ of the a priori possible filtered signals possesses 2^K elements, and a filtering criterion is necessary to decide which of the s_l's is the denoised signal. Of course, the final choice depends on the problem at hand, and for the particular case of gaussian-white and zero-mean noise we propose the CST criterion whose "best" signal s_{CST} verifies

$$CST\left(s_{CST}\right)\simeq p_0 \tag{1}$$

with

$$CST\left(s_l\right)\equiv prob\left(\chi_n^2\leq\frac{\|s-s_l\|^2}{\sigma^2}\right) \tag{2}$$

where χ_n^2 is the Chi-square probability function with n degrees of freedom. The variance, σ^2, of the noise is assumed a priori known, and

$$\|s-s_l\|^2\equiv\sum_{i=1}^{K}\left(s_i-s_{l,i}\right)^2. \tag{3}$$

The probability threshold p_0 in (1) fixes the level of risk accepted that some noise remains in s_{CST}.

In order to show the reader that the choice of a particular filtering criterion is critical and strongly depends on the problem at hand, we consider two general-purpose criteria: the Akaike's Information Criterion (AIC) [*Akaike*, 1965] and the Minimum Description Length (MDL) [*Rissanen*, 1978]. These criteria are often used to choose among a collection of a priori models, like for instance ARMA models in signal processing [*Wax and Kailath*, 1985], with different complexities k. The MDL criterion has been used by *Saito* [1994] in the context of wavelet filtering. When applied to the gaussian case, the AIC and MDL criteria respectively retain the signals s_{AIC} and s_{MDL} such that

$$\begin{aligned}AIC\left(s_{AIC}\right)&=\min\left[AIC\left(s_l\right)\right], &(4)\\ MDL\left(s_{MDL}\right)&=\min\left[MDL\left(s_l\right)\right], &(5)\end{aligned}$$

Paper number 96GL00155
0094-8534/96/96GL-00155$03.00

with

$$AIC(s_l) = \|s - s_l\|^2/\sigma^2 + 2k, \quad (6)$$
$$MDL(s_l) = \|s - s_l\|^2 / (2\sigma^2) + (3/2)\, k \ln K. \quad (7)$$

In our case, k is the number of wavelet coefficients used to produce s_l.

The Denoising Algorithm

The next step is to apply the criteria (1), (4) and (5) to find the output filtered signal into $\mathcal{A}$. Even for short data series, the number of elements in $\mathcal{A}$ is considerable and disables an exhaustive search in the whole set of trial signals. However, the gaussian noise assumption allows a straightforward pre-selection of admissible trial models and limits the search in a subset $\mathcal{B} \subset \mathcal{A}$ with only K elements. Let us define the set $\mathcal{A}_k \subset \mathcal{A}$ formed by the $\binom{K}{k}$ trial signals reconstructed with k wavelet coefficients. Obviously,

$$\bigcup_{k=1}^{K} \mathcal{A}_k = \mathcal{A}. \quad (8)$$

From (2), (6), and (7) we find that, for either criterion, the best candidate s_k belonging to $\mathcal{A}_k$ is such that

$$\|s - s_k\|^2 = \sum_i W_i^2 \quad MINIMUM. \quad (9)$$

where the sum is restricted to the $K - k$ wavelet coefficients W_i discarded to reconstruct s_k. Condition (9) is satisfied if s_k is constructed from the k wavelet coefficients with the largest modulus. Hence, the best signal belonging to each subset $\mathcal{A}_k$ is directly obtained without spanning the whole subset. Equation (8) shows that the initial search performed in the entire set $\mathcal{A}$ may be replaced by a sequential search in the K subsets $\mathcal{A}_k$. In other words, the search is now restricted to the set

$$\mathcal{B} = \{s_k \,;\, k = 1, \cdots, K\} \quad (10)$$

where each element s_k is constructed from the k largest wavelet coefficients. Equation (9) holds because both the wavelet theory and the gaussian statistics rely on the same L^2 norm. If non-gaussian statistics were to be chosen for the noise model, another norm should be used and included in the wavelet transform to produce a fast algorithm similar to the one established in the gaussian case. For instance, an exponential statistics corresponds to the L^1 norm.

Examples

Synthetic Noisy Signals

The synthetic signals, named "Blocks" and "Heavisine" (Figure 1a), are the same as those used by *Donoho and Johnstone* [1994] and have been digitized over 2048 values and contaminated by zero-mean gaussian white noises with $\sigma^2 = 1$ (Figure 1b). For all the following examples, we used the Daubechies' analyzing wavelet possessing 10 vanishing moments [*Daubechies*, 1992]. The AIC-filtered signals still contain a large amount of noise (Figure 1c) which indicates that, for the gaussian statistics, the balance between the fitting and the penalty terms in (6) favors the models with too many degrees of freedom. The MDL- and CST-filtered (with $p_0 \approx 0.5$) signals (Figure 1d,e) are cleaner and we observe limited Gibbs effects in the Blocks signal and several spikes in the Heavisine signal produced by the CST filter. A visual comparison of our results with those obtained by *Donoho and Johnstone* [1994] with a Daubechies' wavelet shows that both the MDL and CST filters almost work like the thresholding used by these authors who found average square errors differing by less than 20% of ours (see Figure 1). Note that smaller errors have been obtained by *Saito* [1994] and by *Donoho and Johnstone* [1994] for the Blocks signal and with the Haar wavelet which, for this particular type of signal, is more efficient than the Daubechies' wavelet used in the present study.

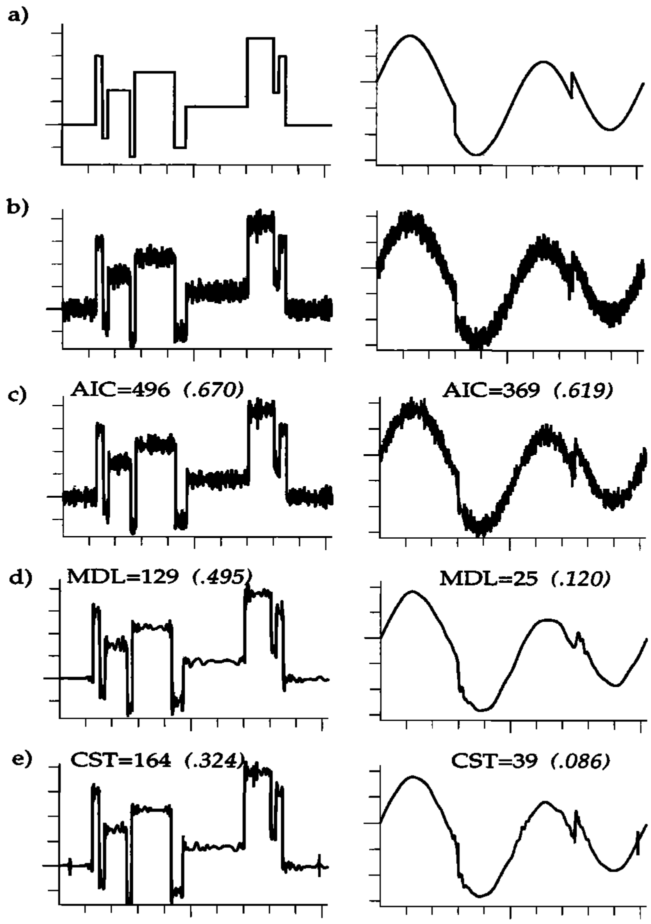

Figure 1. a) Synthetic signals 'blocks' (left) and 'heavisine' (right). b) The same signals contaminated by a gaussian white noise with a unit variance. c) Outputs obtained with the AIC filter, the numbers indicate the number of wavelet coefficients used to produce these reconstructed signals. d) Outputs obtained with the MDL filter. e) Outputs obtained with the CST filter. Numbers in parenthesis are average square errors.

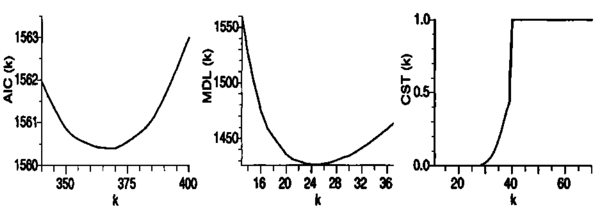

Figure 2. Criteria curves $AIC(s_k)$ (left), $MDL(s_k)$ (middle), and $CST(s_k)$ (right) corresponding to the heavisine signal and plotted for a limited range of k centered on the best value. Notice the flatness of the $AIC(s_k)$ curve and the sharp sigmoidal shape of the $CST(s_k)$ curve.

Figure 2 displays the curves $AIC(s_k)$, $MDL(s_k)$, and $CST(s_k)$ corresponding to the Heavisine signal and plotted for a limited k range centered on the best value found (369, 25, and 39 for AIC, MDL, and CST, respectively). The curves corresponding to the AIC and MDL criteria are almost symmetrical and flat in a wide interval centered on the minimum. Instead, the curve for the CST filter has a sigmoidal shape with a narrow and steep transition zone between the region where the residuals have a very low probability ($\simeq 0$) and becomes a purely gaussian white noise and the region with high probability ($\simeq 1$). This narrow transition zone makes the CST filter unambiguous since a very limited range of possible values for k is associated to the steep edge (Figure 2). As a consequence, the choice of the probability threshold p_0 is not critical and, in most cases, choosing $p_0 > 0.5$ implies a single value for k.

The efficiency of the filters have also be assessed by filtering a pure gaussian white noise with a unit variance. The AIC filter gives a signal reconstructed with a large number (155) of wavelet coefficients. This confirms that the AIC criterion is unable to remove a large part of the noise present in the data. Conversely, the signals obtained with both the MDL and the CST filters are identically null. This agrees with *Saito*'s results [*Saito*, 1994] obtained with the MDL criterion.

Electrical Geophysical Signal

We now illustrate the utility of the method with an application of the MDL and CST filters to thermistance measurements made in a limestone underground quarry [*Morat and Le Mouël*, 1992; *Morat et al.*, 1995]. This signal is strongly non-stationary and possesses abrupt variations (Figure 3). It is particularly interesting to check how the filters are able to account for the abrupt change of regime observed around $t = 2000\ s$. A noise variance $\sigma^2 = 9 \times 10^{-2}\ Ohm^2$ estimated from the high-frequency part of the power spectrum of the entire signal has been used. The MDL- and CST-filtered signals are shown in Figure 3. In accordance with the previous synthetic tests, both filters give qualitatively equivalent results, although the MDL output seems slightly more smoothed than the CST output. A quantitative assessment can be made by examining the Fourier energy spectra of both the initial and filtered signals (Figure 4). Most of the energy of the left half ($t < 2000\ s$) of the initial signal is located in the approximative low-frequency waveband $0 \lesssim f \lesssim 0.03\ Hz$ while the right half ($t > 2000\ s$) has its energy essentially in the $0.015 \lesssim f \lesssim 0.045\ Hz$ waveband (Figure 4). These distinct spectral contents are a consequence of the sudden change of the initial signal around $t = 2000\ s$. The

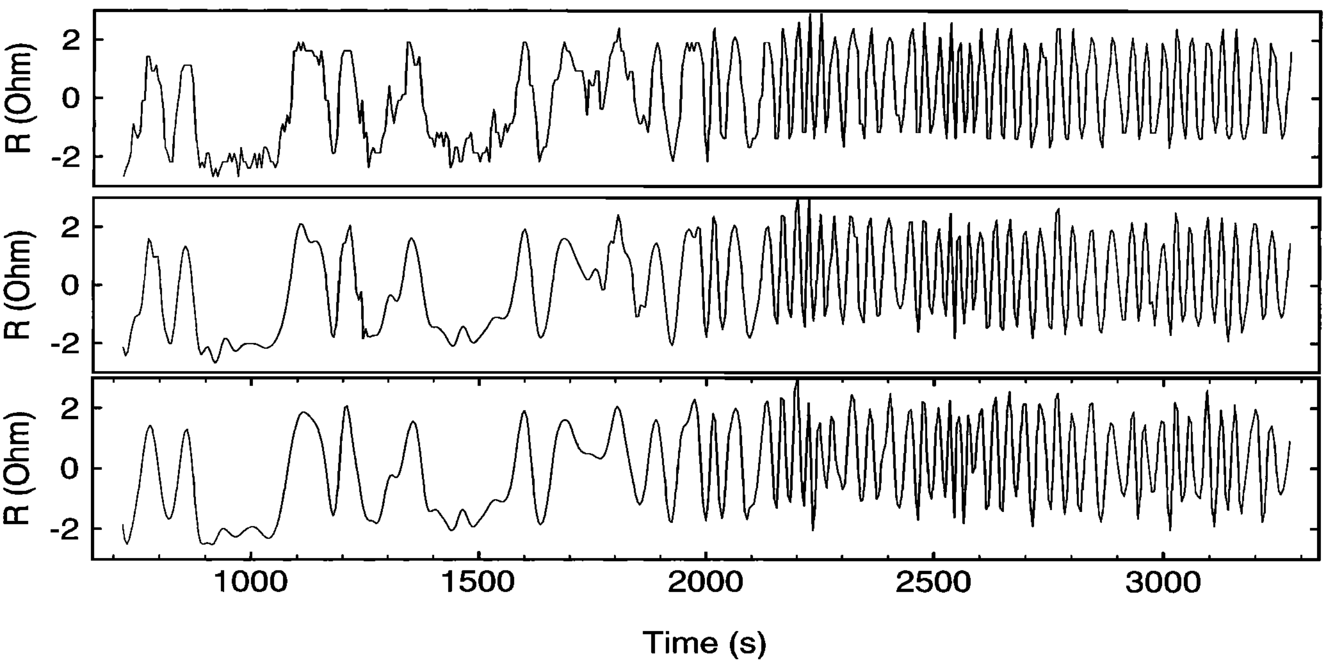

Figure 3. Top: original thermistance data. Middle: CST-filtered signal. Bottom: MDL-filtered signal.

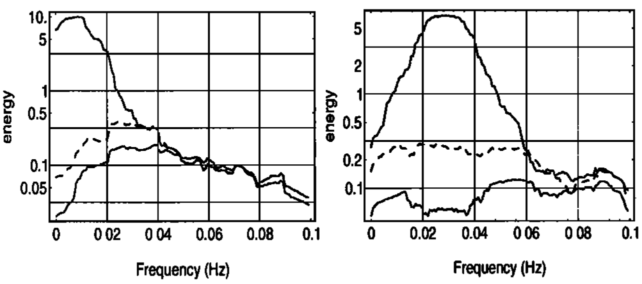

Figure 4. Left: energy spectra of the left halves ($t < 2000\ s$) of the original signal (upper curve), of the noise removed by the CST filter (lower solid curve), and of the noise removed by the MDL filter (dashed curve). Right: same as left part of the figure for the right halves ($t > 2000\ s$) of the signals.

energy spectra of the left halves of the filtered signals show that both the MDL and the CST criteria filtered out the frequencies $f \gtrsim 0.035\ Hz$. We also observe that the CST filter is more efficient to preserve the energy in the low-frequency waveband where most of the information is expected. The energy spectra of the right halves of the filtered signals more clearly illustrate the difference between the filters. In particular, one can see that most of the energy removed by the MDL filter lies in the waveband where the signal-to-noise ratio is high. Conversely, the spectrum of the noise removed by the CST filter indicates that this filter correctly preserved the information waveband. This shows that the CST filter has correctly managed for the non-stationarities of the signal-to-noise ratio: the CST filter roughly acted like a low-pass filter with a cutoff frequency $f_c \simeq 0.03\ Hz$ for the left half of the signal and with $f_c \simeq 0.045\ Hz$ for the right half. The CST filter automatically detects the waveband where the signal-to-noise ratio is high and filters out the frequencies outside this waveband. Such a filtering with a varying and automatically adapted cutoff frequency is impossible by means of classical linear filtering.

Conclusion

Orthogonal wavelet decompositions coupled with the CST criteria is efficient to clean non-stationary signals embedded in gaussian white noise. Contrary to the classical Fourier filtering, the CST filter automatically adapts its cutoff characteristics through a local evaluation of the signal-to-noise ratio. For the gaussian case considered in this letter, the adapted CST criterion is more efficient than the general purpose AIC and MDL criteria. Fast algorithms may be designed as far as the norm used in the scalar product of the wavelet decomposition is compatible with the statistics assumed for the noise.

Acknowledgments. We thank Pierre Morat for kind supplying the thermistance data and helpful comments. We are pleased to acknowledge Alex Grossmann and Jean Morlet for their stimulating discussions. This study was funded by ATP's *Télédétection spatiale* (CNES and CNRS-INSU) and *Valorisation* (CNRS-INSU). The computer programs used to perform this study are available upon request to the authors.

References

Akaike, H., Fitting autoregressive models for prediction, *Ann. Inst. Stat. Math.*, *21*, 261-265, 1965.

Daubechies, I., *Ten Lectures on Wavelets*, *CBMS Ser. Appl. Math.*, Vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

Donoho, D., and I.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, *81*, 425-455, 1994.

Foufoula-Georgiou, E., and P. Kumar, eds., *Wavelets in Geophysics*, 362 pp., Academic, San Diego, Calif., 1994.

Grossmann, A., and J. Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape, *SIAM J. Math.*, *15*, 723-736, 1984.

Holschneider, M., *Wavelets: An analysis tool*, 423 pp., Clarendon, Oxford, 1995.

Mallat, S., Multiresolution approximation and wavelet orthonormal bases of $L^2(\mathbb{R})$, *Trans. Amer. Math. Soc.*, *315*, 69-87, 1989.

Mallat, S., and W.L. Hwang, Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory*, *38*, 617-643, 1992.

Meyer, Y., *Ondelettes et opérateurs*, 383 pp., Hermann, Paris, 1990.

Morat, P., and J.-L. Le Mouël, Signaux électriques engendrés par des variations de contrainte dans des roches poreuses non saturées, *C. R. Acad. Sci., Ser.II*, *315*, 955-963, 1992.

Morat, P., J.-L. Le Mouël, S. Pride, and C. Jaupart, Sur de remarquables oscillations de température, d'humidité, de potentiel électrique observées dans une carrière souterraine, *C. R. Acad. Sci., Ser.II*, *320*, 173-180, 1995.

Rissanen, J., Modeling by shortest data description, *Automatica*, *14*, 465-471, 1978.

Saito, N., Simultaneous noise suppression and signal compression using a library of orthonormal bases and Minimum Description Length criterion, in *Wavelets in Geophysics*, edited by E. Foufoula-Georgiou and P. Kumar, pp. 299-324, Academic, San Diego, Calif., 1994.

Wax, M., and T. Kailath, Detection of signals by information theoretic criteria, *IEEE Trans. Acoust. Speech Signal Processing*, *33*, 387-392, 1985.

D. Gibert, F. Moreau, and G. Saracco, Géosciences Rennes, Université de Rennes 1, Bât. 15 Campus de Beaulieu, 35042 Rennes cedex, France. (e.mail: gibert, moreau, ginet@univ-rennes1.fr)

(received September 27, 1995; revised November 13, 1995; accepted November 27, 1995.)